



Kent Academic Repository

Araujo, Lucas, Fantuzzi, Felipe, Esteves, Pierre M. and Cardozo, Thiago M. (2026)
When carbonyls beat aromaticity: An additive framework for predicting tautomeric preferences. Chemical Science . ISSN 2041-6520.

Downloaded from

<https://kar.kent.ac.uk/115631/> The University of Kent's Academic Repository KAR

The version of record is available from

<https://doi.org/10.1039/D6SC03354G>

This document version

Author's Accepted Manuscript

DOI for this version

Licence for this version

CC BY (Attribution)

Additional information

Versions of research works

Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in **Title of Journal** , Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries

If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our [Take Down policy](https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies) (available from <https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies>).

Chemical Science

Accepted Manuscript

Published on 13 October 2026

[Licensed under CC-BY 4.0](#)



This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.

Journal Name

ARTICLE TYPE

Cite this: DOI: 00.0000/xxxxxxxxxx

When carbonyls beat aromaticity: An additive framework for predicting tautomeric preferences[†]Lucas Araujo,^{a‡} Felipe Fantuzzi,^{*b} Pierre M. Esteves,^{*a} and Thiago M. Cardozo^{*a}

Received Date

Accepted Date

DOI: 00.0000/xxxxxxxxxx

Aromatic structures are often regarded as privileged stabilising motifs, generally retained even when competing structural effects are present. Here, we show that functional-group energetics can overturn this expectation in heteroaromatic tautomerism. Analysis of the 2-hydroxypyridine/2-pyridone equilibrium reveals a near-cancellation between carbonyl stabilisation and the energetic penalty associated with perturbing the aromatic ring framework, accounting for the small experimental energy gap. Extending this approach to pyridines, diazines, and triazines reveals a simple additive pattern: hydroxy/carbonyl exchange favours carbonylated, non-aromatic tautomers, typically requiring two C=O/N–H units to invert the preference, whereas amino/imine and methyl/methylidene exchanges reinforce aromatic tautomers. Spin-coupled generalised valence bond (SCGVB) calculations analysed within the GPF-EP method show that the quasi-classical component favours the aromatic 6c–6e π framework, whereas quantum interference favours localised 2c–2e bonds, especially C=O. This provides an energy-decomposition expression of π -distortivity: without interference, the aromatic tautomers would predominate across all series. Complementary NBO analysis suggests that this trend reflects stronger lactam-type donor–acceptor delocalisation in carbonylated *N*-heteroaromatic tautomers than in the corresponding imine, methylidene, or all-carbon benzenoid keto analogues. The model rationalises benzenoid phenols as boundary cases and predicts tautomeric preferences along the deamination pathway from melamine to cyanuric acid. These results establish a framework for predicting tautomeric preferences and clarify how aromatic-ring stabilisation is balanced against functional-group energetics.

1 Introduction

Tautomerism is a fundamental chemical phenomenon that plays a key role in both organic and inorganic systems.^{1,2} Since the earliest days of structural chemical theory, tautomeric equilibria have shaped our understanding of molecular structure, stability, and reactivity.^{3,4} In organic chemistry, tautomerism influences measurable reactivity trends and synthetic outcomes—affecting, for example, polarographic reductions of carbonyl compounds,⁵ kinetic/thermodynamic regimes in halogenation,⁶ resonance-assisted intramolecular hydrogen bonding in heterocycles,⁷ Schiff

base formation,⁸ azo–hydrazone chromophore switching,⁹ and acid-catalysed carbonyl activation in asymmetric catalysis.¹⁰ It also impacts drug-like scaffolds in solution,¹¹ and extends to crystal phenomena, including self-inhibition of nucleation and growth¹² and tautomerism-induced bending/twisting in biogenic crystals.¹³

In inorganic and organometallic chemistry, tautomerism reshapes coordination and reactivity: classic metal–carbonyl systems display C- versus O-bonding and μ -CO rearrangements,¹⁴ while tautomeric ligand designs can direct Pd-catalysed C–H hydroxylation with O₂.¹⁵ Main-group platforms show analogous patterns, from phosphinylidene P(=O)H $\rightleftharpoons$ P–OH interconversion with substituent-dependent kinetics¹⁶ to boride–borylene ligand tautomerism linked to remote C–C bond formation.¹⁷ In biological and medicinal settings, tautomerism underpins the activity of alkaloids,¹⁸ amino acids,¹⁹ and nucleic acids.^{20–22} Rare tautomers of DNA bases, for example, are directly implicated in mutagenic processes, with transient enol/imino forms hypothesised to mimic Watson–Crick geometry and promote misincorporation.²³ This view is supported by high-resolution crystallography revealing near-isosteric mispairs,²⁴ infrared photodissocia-

^a Instituto de Química, Universidade Federal do Rio de Janeiro, Av. Athos da Silveira Ramos, 149, CT, A-622, Cid. Univ., Rio de Janeiro, RJ 21941-909, Brazil; E-mails: (T.M.C.) thiago@iq.ufrj.br; (P.M.E.) pesteves@iq.ufrj.br.

^b Chemistry and Forensic Science, School of Natural Sciences, University of Kent, Park Wood Rd, Canterbury CT2 7NH, U.K.; E-mail: f.fantuzzi@kent.ac.uk.

[†] Supplementary Information available: Bond enthalpies; energies for all tautomers; complete SCGVB/GPF-EP energy analysis; NBO results; cartesian coordinates. See DOI: 00.0000/00000000.

[‡] Present address: Center for Theoretical Chemistry, Faculty of Chemistry and Biochemistry, Ruhr University Bochum, Universitätsstraße 150, 44801 Bochum, Germany.

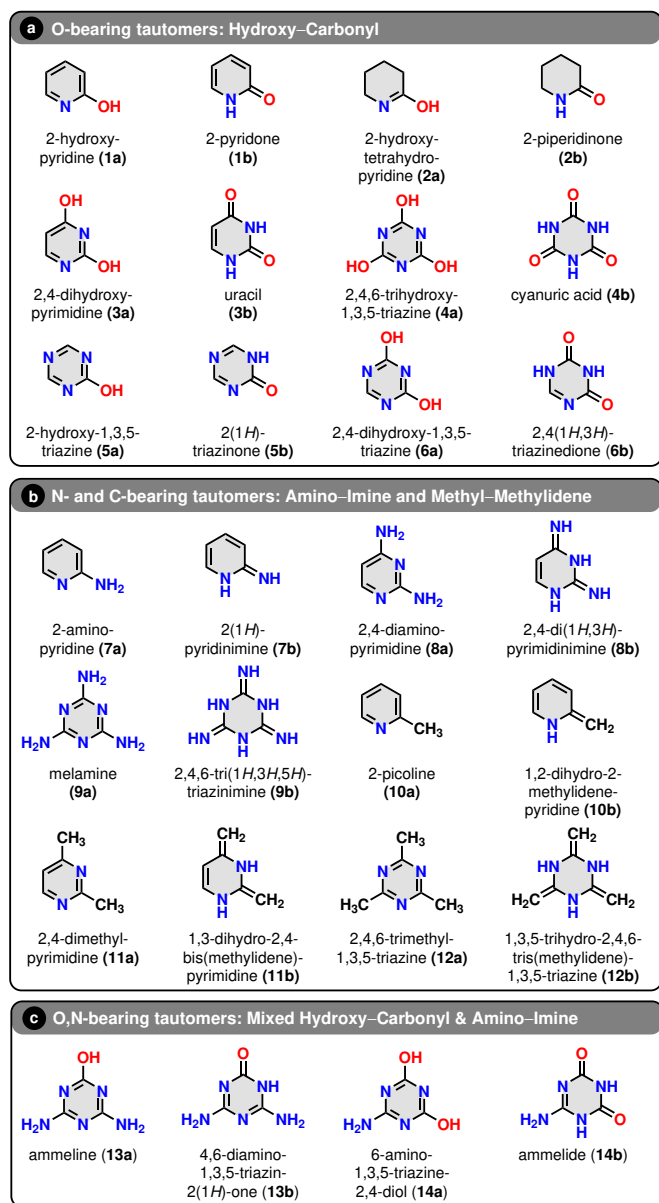


Fig. 1 Heterocyclic systems investigated in this work.

tion evidence for coexisting uracil tautomers,²² and multiscale simulations that map tautomerisation pathways from wobble mismatches to Watson–Crick-like geometries while clarifying environmental contributions.²⁵ More broadly, tautomeric state control modulates binding modes and assay outcomes for drug-like scaffolds and small-molecule therapeutics, including programmes targeting neglected tropical diseases.^{1,26}

One of the main conceptual challenges in tautomerism is its interplay with aromaticity. Aromaticity is among the most widely used yet debated ideas in chemistry, and lacks a universally accepted definition.^{27–29} While Hückel's rule offers a practical rule-of-thumb, its scope is limited.²⁹ In general, equilibria involving an aromatic and a non-aromatic tautomer are expected to favour the aromatic species.^{28,30} Surprisingly, this expectation frequently breaks down: numerous tautomeric equilibria are dominated not by the aromatic species, but by carbonylated

forms.^{20,31,32}

Because tautomerism has long been linked to the concept of aromaticity,³³ numerous theoretical approaches have sought to rationalise tautomeric preferences in these terms. Early work by Katritzky and co-workers introduced empirical resonance-energy schemes to quantify aromatic stabilisation across tautomeric series,^{27,34–36} while Schlegel and co-workers provided benchmark calculations for prototypical systems such as 2-pyridone.³⁷ Complementary orbital-based analyses have also been employed. For example, Hejazi et al.³⁸ used natural bond orbital (NBO)³⁹ calculations to interpret the gas-phase preference for 2-hydroxypyridine (**1a**, Fig. 1a) in terms of steric and electrostatic effects within a Lewis-type framework. More generally, such approaches often invoke donor–acceptor interactions, including $n(\text{N}) \rightarrow \pi^*(\text{C}=\text{O})$ delocalisation in systems such as uracil derivatives.^{40,41} Alternative interpretations have also been proposed, including differing assessments of the roles of steric repulsion, hyperconjugation, and orbital interactions, although the relative importance of these effects remains debated and method-dependent.⁴² Despite this extensive literature, a simple and transferable framework for predicting tautomeric preferences across chemically diverse series remains lacking.

The carbonyl group provides a powerful counterbalance to aromatic stabilisation. The bond dissociation energy of a $\text{C}=\text{O}$ double bond, ca. 178 kcal mol^{−1} (Table S1), exceeds the combined strength of two single $\text{C}–\text{O}$ bonds (ca. 86 kcal mol^{−1} each; ca. 172 kcal mol^{−1} in total), helping to explain the remarkable persistence of carbonyls in drugs, metabolites, and even in fundamental molecules such as carbon dioxide. These tabulated bond-enthalpy increments are not purely local quantities, but are averaged over different chemical structures and therefore implicitly contain stereoelectronic interactions present in those structures. In particular, the enhanced strength of the $\text{C}=\text{O}$ bond has been associated with an anomeric-type $n(\text{O}) \rightarrow \sigma_{\text{C-X}}^*$ interaction, which might contribute to the shortening and strengthening of the $\text{C}=\text{O}$ bond relative to the corresponding $\text{C}=\text{N}$ and $\text{C}=\text{C}$ bonds.^{43,44} In nucleobases such as uracil and thymine, the equilibrium strongly favours the carbonylated tautomer, despite the availability of aromatic alternatives.²¹ The aromaticity of nitrogen-bearing bases has been extensively examined, and it has been suggested that the loss of aromaticity in non-aromatic tautomers may be compensated by stronger local bonding.⁴⁵ This tension between heteroaromaticity and carbonyl stability defines the core problem of many tautomeric equilibria.

A prototypical case is the 2-hydroxypyridine (**1a**)/2-pyridone (**1b**) equilibrium (Fig. 1a), here labelled **R1**, which was first studied by Meyer in 1905.⁴⁶ Experiments consistently show that the aromatic **1a** is preferred in the gas phase, whereas the carbonylated **1b** dominates in condensed phases.^{47–54} These systems have been characterised using mass spectrometry and IR/UV spectroscopy in the gas phase,^{34,55,56} pK_a measurements and NMR in aqueous solution,⁵⁷ and X-ray diffraction and resonant inelastic X-ray scattering (RIXS) in the solid state.^{54,58–60} Another key equilibrium involves 2-aminopyridine (**7a**) and 2(1H)-pyridinimine (**7b**, Fig. 1b), which plays a role in heterocyclic reactivity,^{37,61,62} acid–base chemistry, and even RNA sequenc-

ing through thiolated nucleosides.⁶³ Similar tautomeric preferences are also essential to base pairing and mutagenesis,^{20,55} and molecular aggregation in prebiotic chemistry.⁶⁴

Predicting tautomeric preferences remains challenging because the balance between aromatic stabilisation and carbonyl stability is difficult to anticipate. Here, we introduce a simple qualitative model that captures the principal energetic contributions and enables reliable predictions across a range of heteroaromatic systems. A list of all molecules investigated in this work is shown in Fig. 1.

2 Computational Details

Density functional theory (DFT) calculations were performed using the ORCA 6 software package.⁶⁵ The choice of functional and basis set was validated against DLPNO-CCSD(T)/CBS calculations⁶⁶ using the R1 equilibrium as the benchmark system, given its well-documented sensitivity to electronic-structure treatment. Among all functionals screened with the def2-TZVP basis set, ω B97M-D4rev⁶⁷ delivered the smallest mean absolute deviation (MAD) relative to the coupled-cluster reference, with all range-separated functionals falling within chemical accuracy (Fig. S6a). A subsequent basis-set scan confirmed def2-TZVPP⁶⁸ as the best compromise between cost and accuracy (Fig. S6b), and this combination was therefore adopted throughout. Geometry optimisations and harmonic frequency analyses were carried out at the ω B97M-D4rev/def2-TZVPP level to confirm stationary points as minima (no imaginary frequencies) and to obtain thermochemical functions within the ideal-gas, rigid-rotor harmonic-oscillator (RRHO) approximation. Unless stated otherwise, reported thermodynamic quantities are derived from these frequency calculations.

Valence bond calculations were performed using VB2000 interfaced with the GAMESS-US package.^{69,70} Chemical-bond analyses were carried out using the generalised product function energy partitioning (GPF-EP) method,⁷¹ as implemented in VB2000. This framework is particularly suited to the present problem because it provides direct access to the energetic contributions associated with self-consistently optimised non-orthogonal one-electron orbitals and the interference terms arising from their interactions.

GPF-EP applies the Ruedenberg density partitioning⁷² to a generalised product function (GPF),⁷³ in which the N -electron wavefunction is expressed as an antisymmetrised product of strongly orthogonal bond functions:

$$\Psi_{\text{GPF}} = \mathcal{A} \{ \Psi^{(1)}, \Psi^{(2)}, \dots, \Psi^{(G)} \} \quad (1)$$

In Ruedenberg's formulation, the accumulation of electron density between bonded nuclei is a quantum-mechanical interference effect, analogous in origin to the density redistribution observed in the double-slit experiment.⁷² This density relocation by interference provides a quantum-mechanical basis for chemical bond formation and is illustrated schematically in Fig. 2. Within the GPF-EP framework, the total energy can therefore be separated into contributions with a quasi-classical interpretation, such as electrostatics and charge polarisation, and contributions arising

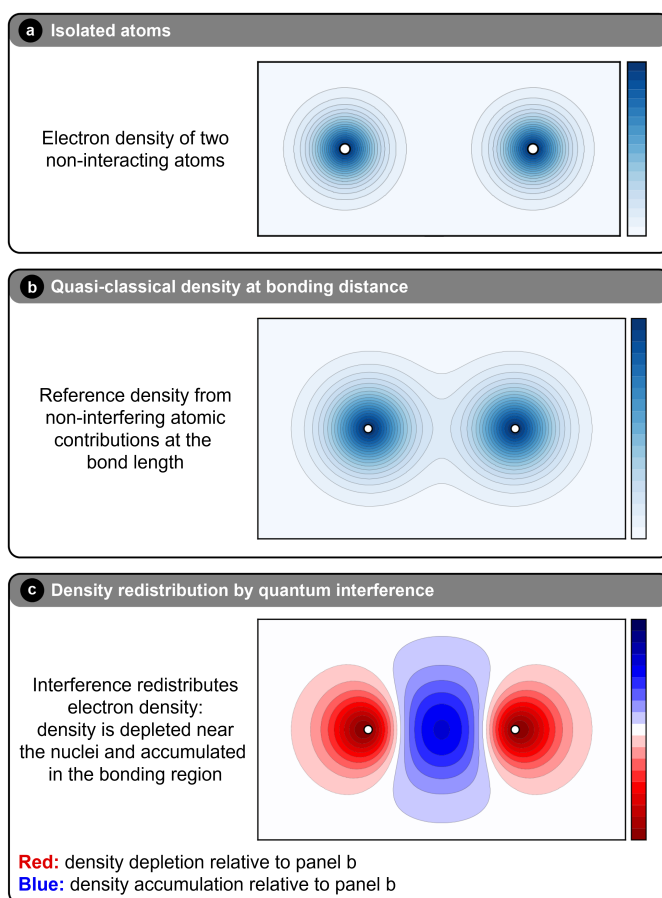


Fig. 2 Schematic representation of the Ruedenberg density partitioning underlying the GPF-EP analysis. (a) Electron density of two isolated, non-interacting atoms. (b) Quasi-classical reference density obtained by placing the same atomic contributions at the bonding distance, without orbital interference. (c) Interference density, defined as the difference between the total density of the interacting system and the quasi-classical reference density in panel b. Red regions indicate density depletion near the nuclei, whereas blue regions indicate density accumulation in the interatomic bonding region.

from orbital interference. A useful feature of the method is that this partition can be applied to selected electron groups, such as individual 2c–2e bonds or larger delocalised groups, including the six π electrons of benzene. A complete derivation is given in Section S2 of the Supporting Information; here, we summarise the key elements.

For each electron group μ , the one-electron density is exactly decomposed into a quasi-classical component, corresponding to the sum of squared individual orbital densities, and an interference component, arising from orbital non-orthogonality:

$$\rho^{(\mu)} = \rho_{\text{QC}}^{(\mu)} + \rho_{\text{INT}}^{(\mu)} \quad (2)$$

The interference density integrates to zero, but redistributes electron density within the molecule, most notably from the nuclear centres into bonding regions.

This density decomposition propagates through the pair density

and, consequently, through the total energy:

$$E[\text{TOT}] = E[\text{QC}] + E[\text{INT}] \quad (3)$$

where $E[\text{QC}]$ is the exchange-corrected quasi-classical energy, including independent-orbital density terms, Coulomb interactions, and inter-group Pauli exchange. Although exchange is quantum-mechanical in origin, it is included in $E[\text{QC}]$ because it arises from antisymmetrisation between electron groups, rather than from the intra-group density redistribution caused by orbital interference. The $E[\text{INT}]$ term is the total interference energy, associated with the energetic effect of density redistribution caused by constructive and destructive interference between non-orthogonal orbitals. In chemical terms, $E[\text{INT}]$ captures the covalent stabilisation associated with bond formation, whereas $E[\text{QC}]$ contains contributions that can be interpreted without invoking orbital interference, such as charge polarisation, electrostatic interactions and repulsion between electron pairs, among others.

All GPF calculations employed a two-group partition. One group comprised the six π electrons of the heteroaromatic ring, corresponding to the Hückel sextet in pyridines, diazines, and triazines. For the corresponding non-aromatic tautomers, the same group comprised the six electrons occupying the formal π framework derived from the ring, i.e., the localised pairwise π bonds that replace the aromatic sextet. This group was treated with SCGVB, allowing all spin couplings among the six electrons. The second group contained the remaining electrons, including the σ framework, lone pairs, and exocyclic substituents, and was treated at the Hartree-Fock level. The interference analysis was carried out for the first group only, because previous work showed that energy changes associated with distortions of the σ framework are predominantly quasi-classical.⁷⁴ Preliminary tests using alternative partitions and larger active groups indicated that this scheme provides a good compromise between wavefunction quality and the convergence demands of non-orthogonal SCGVB optimisation. The cc-pVDZ basis set was used throughout.

Before discussing the results, we briefly state the modern valence-bond picture of bonding in benzene-like aromatics that underpins our analysis. In this framework, the π system is described as a six-centre, six-electron (6c–6e) delocalised bond, an SCGVB description first articulated by Cooper *et al.*⁷⁵ and subsequently refined in later work.^{74,76} From the interference perspective, adjacent p -like orbitals in the π framework exhibit net constructive interference around the ring, as reflected in both the pairwise interference energy terms and the corresponding interference density. The systems considered here are conventional cases in which the aromatic or non-aromatic character of the relevant tautomers is well established: the **a**-tautomers retain a cyclic six- π -electron framework, whereas the corresponding **b**-tautomers contain localised exocyclic double bonds and no longer retain the same delocalised ring pattern. Throughout this work, we therefore use the presence of a 6c–6e π bond with net constructive interference as an operational SCGVB/GPF-EP criterion to identify aromatic bonding patterns within these well-defined systems. We do not attempt to define or quantify aromaticity more broadly, given the continuing debate surrounding

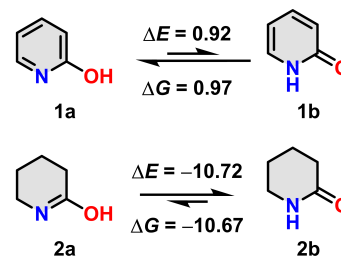


Fig. 3 Energetics of the 2-hydroxypyridine (**1a**) $\rightleftharpoons$ 2-pyridone (**1b**) equilibrium and its saturated reference, 2-hydroxytetrahydropyridine (**2a**) $\rightleftharpoons$ 2-piperidinone (**2b**). Energies (in kcal mol⁻¹) at the ω B97M-D4rev/def2-TZVPP level.

the term.²⁹

3 Results and Discussion

3.1 The 2-hydroxypyridine/2-pyridone equilibrium and related series

Following Katritzky's concept of group transferability,³⁴ we hypothesise that tautomeric preferences can be rationalised by additive contributions from the ring and from the exchanging functional group. Focusing on the 2-hydroxypyridine (**1a**) $\rightleftharpoons$ 2-pyridone (**1b**) equilibrium (Fig. 3, top), we define the energy difference as:

$$\Delta E = E_{1b} - E_{1a} \approx \Delta E_R + \Delta E_{FG} \quad (4)$$

where ΔE_R captures changes in the ring π -framework (aromaticity gain/loss, ring strain, and polarisation) and ΔE_{FG} collects local bond-energy changes within the interconverted moiety, implicitly including differences in local effects such as electrostatic interactions and stereoelectronic effects involving the functional groups. For **1a** $\rightarrow$ **1b**, this moiety is (O–H, C=N) $\rightarrow$ (N–H, C=O); a rough estimate from tabulated bond enthalpies (Table S1) gives $\Delta E_{FG} \approx -13.17$ kcal mol⁻¹, reflecting the strong C=O bond. The DFT gap is $\Delta E(\mathbf{1a} \rightarrow \mathbf{1b}) = 0.92$ kcal mol⁻¹ (Table S2, entry **1a** $\rightarrow$ **1b**), in line with experiment (0.77 kcal mol⁻¹).⁴⁷ Hence, ΔE_R is on the order of +13 kcal mol⁻¹; specifically, $\Delta E_R(\mathbf{1a} \rightarrow \mathbf{1b}) = +14.09$ or +13.94 kcal mol⁻¹ when using the DFT or experimental $\Delta E(\mathbf{1a} \rightarrow \mathbf{1b})$, respectively—nearly compensating the carbonyl stabilisation. We use this simple decomposition to rationalise trends across related heteroaromatic series and to formulate an additive rule for substituent effects.

An independent estimate of $\Delta E_R(\mathbf{1a} \rightarrow \mathbf{1b})$ is obtained by comparing the pyridine pair with its non-aromatic analogue, the δ -lactim/ δ -lactam equilibrium in 2-hydroxytetrahydropyridine (**2a**) to 2-piperidinone (**2b**; Fig. 3, bottom). In this saturated reference, the ring contribution is minimised, so the observed gap approximates the functional-group term, $\Delta E_{FG} \approx \Delta E_{\text{sat}}$. Accordingly:

$$\Delta E_R \approx \Delta E_{\text{aromatic}} - \Delta E_{\text{sat}} \quad (5)$$

with $\Delta E_{\text{aromatic}} = \Delta E(\mathbf{1a} \rightarrow \mathbf{1b})$ and $\Delta E_{\text{sat}} = \Delta E(\mathbf{2a} \rightarrow \mathbf{2b})$. Using the tabulated values (Table S2, entries **1a** $\rightarrow$ **1b** and **2a** $\rightarrow$ **2b**) gives $\Delta E_R(\mathbf{1a} \rightarrow \mathbf{1b}) \approx +11.64$ kcal mol⁻¹, close to the bond-enthalpy-

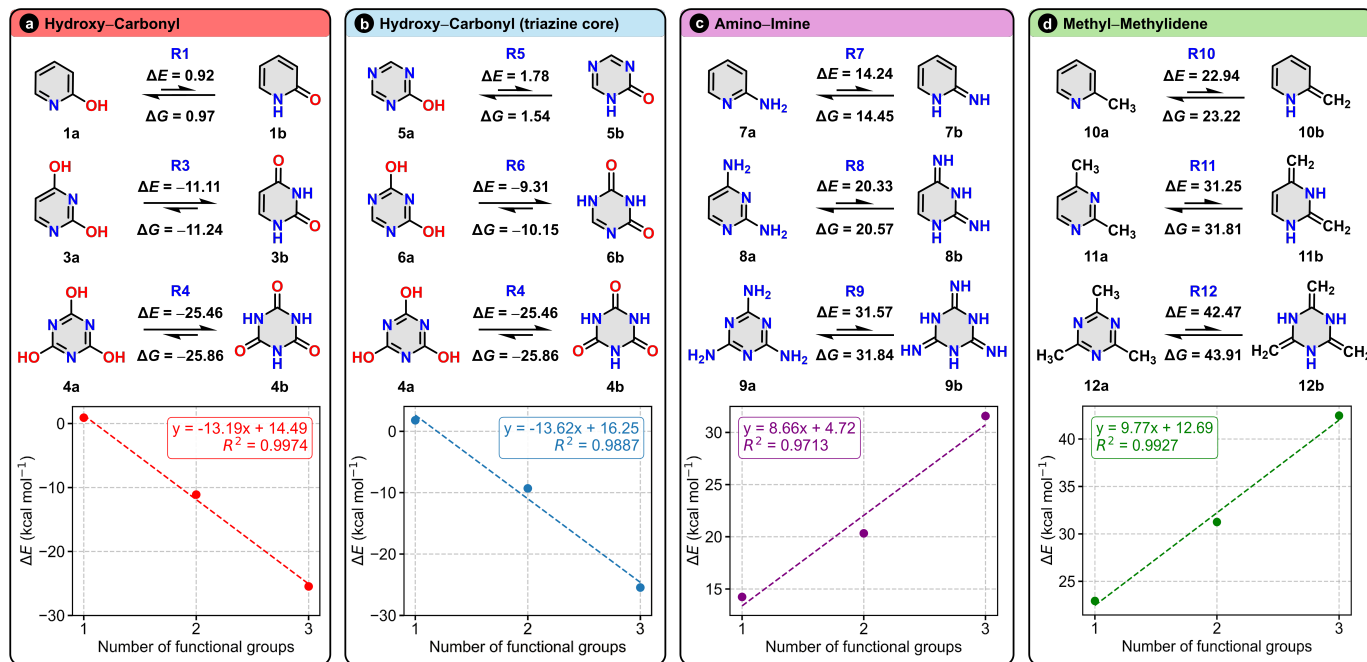


Fig. 4 Relative energies (in kcal mol⁻¹) for homologous tautomeric sets as the number and type of exchanging groups increase. (a) Hydroxy-carbonyl pairs **1a/1b**, **3a/3b**, **4a/4b**; (b) Hydroxy-carbonyl pairs on a fixed triazine core **5a/5b**, **6a/6b**, **4a/4b**; (c) Amino-imine pairs **7a/7b**, **8a/8b**, **9a/9b**; (d) Methyl-methylidene pairs **10a/10b**, **11a/11b**, **12a/12b**. In each set, the upper part shows calculated ΔE and ΔG with structures; the lower plots ΔE versus the count of exchanging groups. Reactions are labelled RX, with X matching the tautomeric pair. Energies (in kcal mol⁻¹) were obtained at the ω B97M-D4rev/def2-TZVPP level.

based estimate. While part of $\Delta E_R(\mathbf{1a} \rightarrow \mathbf{1b})$ can be ascribed to loss of the aromatic π pattern, residual ring-structure effects (e.g., polarisation, geometry, hyperconjugation) also contribute. In net, the stabilisation from carbonyl formation is almost exactly offset by ring-framework penalties, yielding the observed small $\Delta E(\mathbf{1a} \rightarrow \mathbf{1b})$.

We next analyse the uracil tautomeric pair: 2,4-dihydropyrimidine (**3a**, lactim) and uracil (**3b**, lactam). Assuming additivity for the functional-group term, introducing an additional (C=O, N-H) unit relative to the **1a/1b** pair contributes approximately -13.17 kcal mol⁻¹ to ΔE . With the convention $\Delta E = E_{3b} - E_{3a}$, this gives $\Delta E(\mathbf{3a} \rightarrow \mathbf{3b}) \approx \Delta E(\mathbf{1a} \rightarrow \mathbf{1b}) - 13.17 = -12.25$ kcal mol⁻¹, in excellent agreement with the DFT value, $\Delta E(\mathbf{3a/3b}) = -11.11$ kcal mol⁻¹ (Table S2, entry **3a** → **3b**). Despite the simplicity of the additivity model, this consistency supports using the ring/functional-group decomposition to anticipate tautomeric preferences across related systems.

Although the present analysis focuses on gas-phase tautomeric preferences, the reader may be interested in the effects of solvation on tautomeric equilibria. It has long been established that, for equilibria such as **R1**, solvents with high dielectric constants favour the carbonylated form, whereas solvents with low dielectric constants tend to maintain a more balanced distribution of tautomers (see ref.⁷⁷). Experimentally, the dominance of the lactam form **1b** in aqueous solution prompted us to examine the effect of implicit solvation on the prototypical **R1** equilibrium. As shown in Fig. S2a, inclusion of water via the CPCM⁷⁸ and SMD⁷⁹ continuum models reverses the slight gas-

phase preference ($\Delta E = +0.92$ kcal mol⁻¹ in favour of **1a**), yielding $\Delta E \approx -4$ kcal mol⁻¹ in favour of the carbonylated tautomer **1b** in both solvation models. This result is consistent with experimental data obtained in aqueous solution.^{80,81} We attribute this reversal to differential electrostatic stabilisation: the larger dipole moment and enhanced hydrogen-bond-accepting capacity of the lactam form are better accommodated in a polar protic environment. This changes the effective magnitude of ΔE_R while leaving ΔE_{FG} comparatively unchanged (vide Fig. S2b). Consequently, the functional-group term changes only modestly in polar solvents, whereas the aromatic ring penalty is affected by dielectric screening and polarisation effects.

A comprehensive solvation analysis of all tautomeric pairs is beyond the scope of the present work. We emphasise that the predictive model developed here primarily describes gas-phase tautomeric preferences, and its direct application to solution-phase equilibria requires explicit treatment of solvation effects. Nevertheless, these results confirm that the ring/functional-group decomposition framework remains qualitatively valid across phases: solvation changes the effective ring/polarisation contribution in favour of the carbonylated tautomer, while the functional-group term remains comparatively transferable.

Finally, the additive framework developed here is most directly applicable to six-membered N-containing heteroaromatic rings, for which ΔE_R is well-defined and ΔE_{FG} can be approximated using transferable bond-enthalpy increments. Extension to *n*-membered heteroaromatic rings would require reparameterisation of ΔE_R due to differences in electronic structure and conformational effects. Moreover, the transferability of ΔE_{FG} is ex-

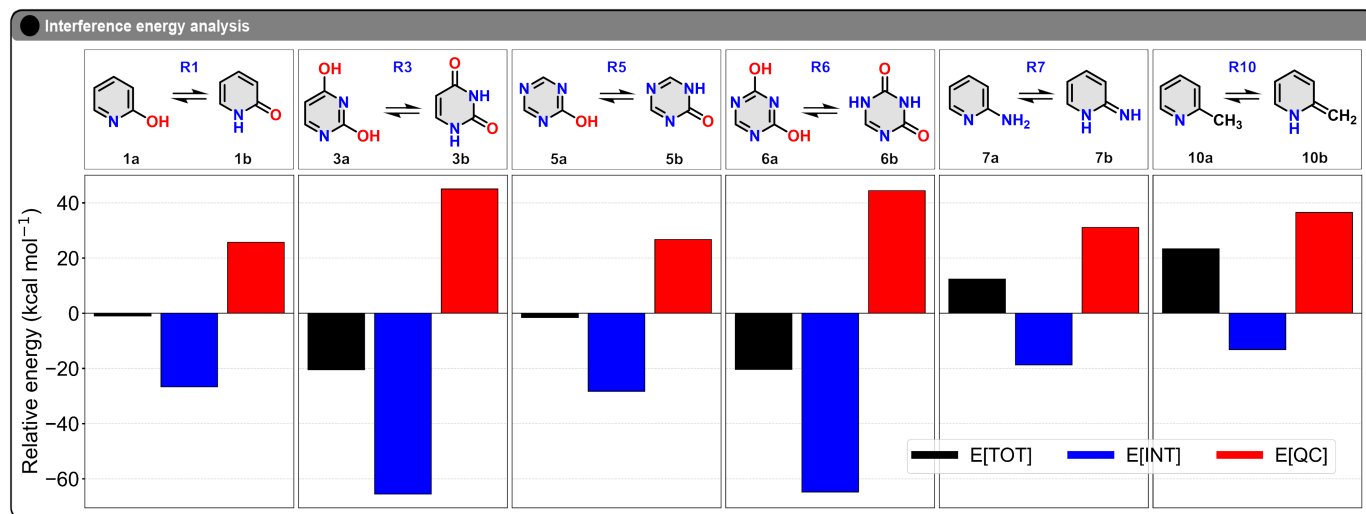


Fig. 5 Interference-energy analysis of selected tautomeric equilibria within the SCGVB/GPF-EP framework. The total energy is partitioned into quasi-classical, $E[QC]$, and interference, $E[INT]$, contributions. In the absence of the interference contribution, the aromatic tautomers would prevail across the series. Numerical values for the total decomposition are provided in Tables S4–S5, while the corresponding interference-density plots and pairwise p -orbital interference terms are shown in Figs. S3–S4 and Table S6.

pected to be greatest when functional groups are directly conjugated to, or in close electronic communication with, the aromatic ring. Nevertheless, the underlying principle that cumulative formation of multiple C=O/N–H units can overcome aromatic stabilisation and shift equilibria towards carbonylated tautomers is expected to remain broadly applicable.

3.2 Carbonyls vs. 6c–6e bonds

Fig. 4 summarises the relative energies for homologous tautomeric sets as the number and type of exchanging groups increase. The figure is organised as follows: (a) hydroxy–carbonyl pairs **1a/1b**, **3a/3b**, **4a/4b**, increasing the number of OH groups in the enolic tautomer from 1 to 3 while simultaneously moving from pyridine to triazine; (b) hydroxy–carbonyl pairs on a fixed triazine core with a stepwise increase in OH groups (pairs **5a/5b**, **6a/6b**, **4a/4b**); (c) amino–imine pairs traversing the heteroaromatic scaffold from pyridine to triazine (pairs **7a/7b**, **8a/8b**, **9a/9b**); and (d) methyl–methylidene pairs, increasing from one to three CH_3 substituents (pairs **10a/10b**, **11a/11b**, **12a/12b**), while likewise progressing from pyridine to triazine. In each case, the upper portion shows the calculated ΔE and ΔG for the individual tautomeric transformations together with the corresponding structures, and the lower portion plots ΔE against the number of exchanging functional groups. Each transformation is labelled as **RX**, where X matches the label of the corresponding tautomeric pair. While each series is limited, chemical consistency of the slopes across structurally related series provides qualitative support for the additive model. We present these correlations as illustrations of the additivity principle rather than statistically validated quantitative predictions.

When only a single tautomeric group is present (e.g., **R1**, **R5**), the aromatic tautomer is only marginally favoured; the small $\Delta E/\Delta G$ reflects near-cancellation between the local stabilisation gained on forming a C=O/N–H unit and the penalty associ-

ated with perturbing the ring π -framework. Introducing additional O-bearing groups (hydroxy–carbonyl) drives a monotonic shift towards the carbonylated structures, both along the scaffold change from pyridine to triazine (**R1**→**R3**→**R4**) and on a fixed triazine core (**R5**→**R6**→**R4**). In the lower panels, ΔE varies approximately linearly with the number of (C=O,N–H) units, consistent with an additive model. Linear fits give slopes of $-13.2 \text{ kcal mol}^{-1}$ for (a) and $-13.6 \text{ kcal mol}^{-1}$ for (b), an average of $-13.4 \text{ kcal mol}^{-1}$ per C=O/N–H formed. Accordingly, inversion of the preference typically requires two such units. The corresponding ΔG values follow the same trend at ambient temperature, indicating that entropic contributions are modest relative to the dominant enthalpic term.

By contrast, for N- and C-bearing pairs (amino–imine; methyl–methylidene), the aromatic tautomer is reinforced: exocyclic C=N or C=CH₂ formation is disfavoured relative to retention of an NH₂ or CH₃ donor on an aromatic ring, which better supports the ring π -system and avoids the additional polarisation/strain associated with extra exocyclic multiple bonds. Quantitatively, the ΔE –count correlations display positive slopes of $+8.7$ and $+9.8 \text{ kcal mol}^{-1}$ for (c) and (d), respectively. Across these sets, ΔE spans 14.24 – $31.57 \text{ kcal mol}^{-1}$ for (c) and 22.94 – $42.47 \text{ kcal mol}^{-1}$ for (d), and the corresponding ΔG values closely track ΔE at ambient temperature, indicating minimal entropic modulation of the overall trend. As a result, the non-aromatic alternatives remain higher in energy throughout these series. A related boundary case is provided by benzenoid phenols: carbonyl formation stabilises the keto forms, but the net gas-phase ΔE remains in favour of the enol across phenol, resorcinol, and phloroglucinol; the enol–keto gap decreases as the number of OH groups increases but does not invert under these conditions (Table S3 and Fig. S1), as discussed further below.⁸²

Building on these additive trends, we next examine how the energy balance is resolved within the SCGVB/GPF-EP frame-

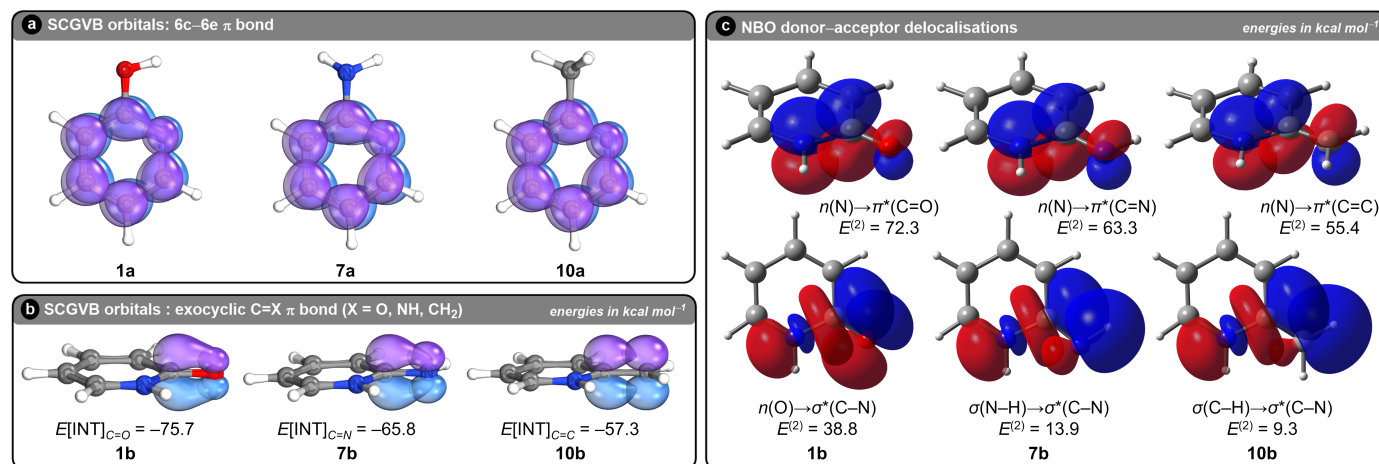


Fig. 6 SCGVb/GPF-EP and NBO analysis of the singly substituted equilibria R1, R7, and R10. (a) SCGVb orbital representation of the aromatic 6c–6e π bond in 1a, 7a, and 10a. (b) Localised exocyclic C=X π bonds in 1b, 7b, and 10b (X = O, NH, CH₂), with the corresponding pairwise interference energies, $E[\text{INT}]_{\text{C=X}}$, in kcal mol⁻¹. Negative values denote stabilising interference contributions. (c) Selected NBO donor–acceptor delocalisations in the non-aromatic tautomers. The upper row shows resonance-type $n(\text{N}) \rightarrow \pi^*(\text{C=X})$ interactions associated with the exocyclic C=X unit, whereas the lower row shows the principal $\sigma^*(\text{C–N})$ -acceptor interactions used in the stereoelectronic comparison. Second-order perturbation energies, $E^{(2)}$, are given in kcal mol⁻¹ as positive stabilisation energies.

work. Fig. 5 decomposes the total energy ($E[\text{TOT}]$) into quasi-classical ($E[\text{QC}]$) and interference ($E[\text{INT}]$) contributions for representative tautomerisation reactions. Across all cases analysed, the two components act in opposition: the $E[\text{QC}]$ term consistently favours an aromatic arrangement of the ring π system, whereas the $E[\text{INT}]$ term favours localisation into stronger 2c–2e bonds, including carbonyl and other exocyclic double bonds. In other words, in the hypothetical absence of quantum interference, all series would prefer the aromatic a-forms; the preference for carbonyl-bearing tautomers, where observed, is enabled by the stabilising $E[\text{INT}]$ contribution. This picture accords with π -distortivity^{83,84}: interference promotes double-bond formation and bond-length alternation, while $E[\text{QC}]$ resists localisation. The corresponding pairwise p -orbital interference matrices and interference-density plots, reported in the Supporting Information, show this reorganisation explicitly: aromatic tautomers retain a delocalised 6c–6e signature, whereas carbonylated tautomers display localised patterns consistent with bond alternation. It also echoes our previous SCGVb/GPF-EP analysis of benzene,⁷⁴ wherein interference stabilises Kekulé-type structures through the formation of three strong 2c–2e bonds, consistent with experimental indications of π -driven distortivity.^{85,86} Large shifts in tautomer preference arise when the interference gain from forming localised bonds (notably C=O) outweighs the $E[\text{QC}]$ bias towards the aromatic structure. Small residual discrepancies in sign or magnitude in cases with small energy differences reflect modest missing dynamic correlation in the SCGVb description without altering these qualitative conclusions.

To understand why carbonyl formation, but not imine or methylenide formation, can overturn the aromatic preference, we focus on the singly substituted equilibria **R1**, **R7**, and **R10** (Fig. 6). These three systems undergo the same formal conversion from an aromatic **a** tautomer to a non-aromatic **b** tautomer, but differ in the exocyclic unit formed: C=O, C=NH, or

C=CH₂. The SCGVb orbitals of **1a**, **7a**, and **10a** show the six non-orthogonal orbitals that combine to form the aromatic 6c–6e π bond (Fig. 6a). The close similarity among these orbital patterns indicates that the three aromatic tautomers share a comparable electronic structure before tautomerisation. We note that SCGVb orbitals are allowed to polarise, thereby incorporating simple ionic-like contributions into the wavefunction,⁸⁷ which improves the quantitative description without changing the qualitative bonding assignment.

In the corresponding **b** tautomers, this 6c–6e arrangement is replaced by a localised exocyclic C=X π bond (Fig. 6b). The pairwise interference energies for these exocyclic bonds become progressively less stabilising from **1b** to **7b** and **10b**, with $E[\text{INT}]_{\text{C=X}}$ values of –75.7, –65.8, and –57.3 kcal mol⁻¹, respectively. This local trend parallels the total $\Delta E[\text{INT}]$ values for the corresponding tautomerisations, which decrease in stabilising magnitude from **R1** to **R7** and **R10** (–26.7, –21.1, and –14.3 kcal mol⁻¹, respectively). Carbonyl formation therefore gives the largest interference stabilisation among the three exocyclic units, although the final tautomeric preference still depends on its balance with the quasi-classical penalty. The other side of the balance is the quasi-classical term. In all three reactions, $E[\text{QC}]$ favours the aromatic tautomer, with $\Delta E[\text{QC}]$ values of +25.7, +34.6, and +40.3 kcal mol⁻¹ for **R1**, **R7**, and **R10**, respectively. The slight preference for **1b** therefore arises only because the interference stabilisation is large enough to overcome this quasi-classical penalty. In **R7** and **R10**, the interference term is weaker and the quasi-classical penalty is larger, so the aromatic tautomers remain preferred.

We next used NBO analysis to relate these trends to a more conventional stereoelectronic description of tautomerism. Such analyses are often framed in terms of hyperconjugation and donor–acceptor delocalisation, in which occupied lone-pair, σ , or π donor orbitals interact with antibonding acceptor orbitals. This

stereoelectronic language has been emphasised by Alabugin and co-workers in the context of hyperconjugation, anomeric effects, donor–acceptor orbital interactions, the stereoelectronic role of oxygen, and the stabilisation of carbonyl, amide, and urea motifs.^{40,43,44,88} The NBO results provide complementary explanation for why carbonyl formation has different consequences in *N*-heteroaromatic and all-carbon benzenoid systems. In the *N*-heteroaromatic series, carbonylated tautomers become preferred once two or more C=O/N–H units are present. In the benzenoid phenol series, carbonyl formation also reduces progressively the enol–keto energy gap but is not sufficient to invert the tautomeric preference under the gas-phase conditions considered here (Table S3 and Fig. S1).

The selected NBO interactions support two complementary comparisons (Fig. 6c and Table S7). First, the resonance-type $n(\text{N}) \rightarrow \pi^*(\text{C}=\text{X})$ interactions in the non-aromatic *N*-heteroaromatic tautomers decrease from the lactam-type carbonyl tautomer **1b** ($E^{(2)} = 72.3 \text{ kcal mol}^{-1}$) to the imine tautomer **7b** ($E^{(2)} = 63.3 \text{ kcal mol}^{-1}$) and the methyldene tautomer **10b** ($E^{(2)} = 55.4 \text{ kcal mol}^{-1}$). Second, the all-carbon benzenoid keto tautomer provides two useful reference points (Fig. S5 and Table S7). Its dominant conjugative interaction is the enone-like $\pi(\text{C}=\text{C}) \rightarrow \pi^*(\text{C}=\text{O})$ donation ($E^{(2)} = 29.1 \text{ kcal mol}^{-1}$), which is much smaller than the lactam-type $n(\text{N}) \rightarrow \pi^*(\text{C}=\text{O})$ interaction in **1b**. A more site-specific comparison is obtained by replacing the ring *N*-lone-pair donor in **1b** with the benzenoid CH_2 unit present in the keto tautomer of phenol. The corresponding $\sigma(\text{C}–\text{H}) \rightarrow \pi^*(\text{C}=\text{O})$ hyperconjugative donations are only 7.0 and 7.2 kcal mol^{-1} , or 14.2 kcal mol^{-1} combined. Thus, replacing the lactam *N*-lone-pair donor by CH_2 removes a strong resonance interaction, providing a stereoelectronic reason why carbonyl formation is less able to overcome aromaticity in the benzenoid series.

The selected $\sigma^*(\text{C}–\text{N})$ -acceptor interactions shown in Fig. 6c give a related but less dramatic comparison. Within the non-aromatic **b**-tautomers, the carbonylated *N*-heteroaromatic tautomer **1b** has the strongest interaction, $n(\text{O}) \rightarrow \sigma^*(\text{C}–\text{N})$ ($E^{(2)} = 38.8 \text{ kcal mol}^{-1}$). The corresponding interactions are weaker in the imine and methyldene tautomers: $\sigma(\text{N}–\text{H}) \rightarrow \sigma^*(\text{C}–\text{N})$ in **7b** ($E^{(2)} = 13.9 \text{ kcal mol}^{-1}$) and $\sigma(\text{C}–\text{H}) \rightarrow \sigma^*(\text{C}–\text{N})$ in **10b** ($E^{(2)} = 9.3 \text{ kcal mol}^{-1}$). The carbonylated tautomer **1b** also has a stronger $n(\text{O}) \rightarrow \sigma^*(\text{C}–\text{N})$ interaction than the analogous $n(\text{O}) \rightarrow \sigma^*(\text{C}–\text{C}_{\text{sp}^3})$ interaction in the all-carbon benzenoid reference ($E^{(2)} = 29.3 \text{ kcal mol}^{-1}$). The imine case is conformationally sensitive. The higher-energy imine isomer **7b'** is included in Table S7 because it allows the more directly comparable $n(\text{N}) \rightarrow \sigma^*(\text{C}–\text{N})$ interaction; this term has $E^{(2)} = 21.8 \text{ kcal mol}^{-1}$, still substantially smaller than the corresponding carbonyl interaction in **1b**.

A differential comparison between the aromatic **a** and non-aromatic **b** tautomers gives the same qualitative trend. For the selected $\sigma^*(\text{C}–\text{N})$ -acceptor interactions associated with the tautomerising unit, the increase in $E^{(2)}$ is largest for **1a**→**1b** (+30.9 kcal mol^{-1}), smaller for **7a**→**7b** (+8.2 kcal mol^{-1}), and smallest for **10a**→**10b** (+4.4 kcal mol^{-1}). Because NBO $E^{(2)}$ values are perturbative donor–acceptor descriptors rather than ther-

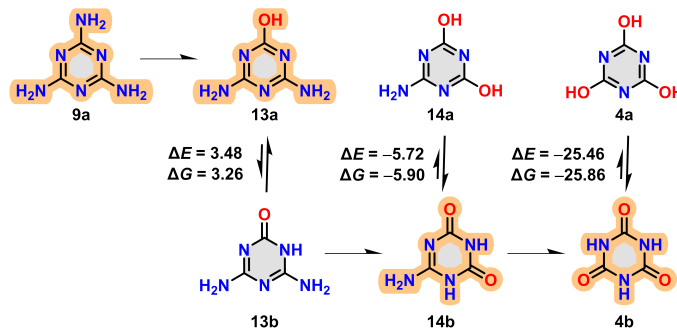


Fig. 7 Successive deamination of melamine on a 1,3,5-triazine core, showing dominant tautomers and computed $\Delta E/\Delta G$ for each tautomerisation and deamination step. Amino/imine substituents stabilise the aromatic tautomer, whereas hydroxy/carbonyl units favour the localised, non-aromatic (lactam) tautomer; two C=O/N–H units typically overturn the aromatic preference. Accordingly, melamine (**9a**) and ammeline (**13a**) are aromatic; ammelide (**14b**) and cyanuric acid (**4b**) are non-aromatic. Energies (in kcal mol^{-1}) at the $\omega\text{B97M-D4rev/def2-TZVPP}$ level.

modynamic energy components, this comparison is used qualitatively; nevertheless, it supports the same ordering inferred from the SCGVB/GPF-EP analysis.

This stereoelectronic picture complements the SCGVB/GPF-EP energy decomposition. Within the SCGVB/GPF-EP framework, carbonyl formation gives strong local interference stabilisation through formation of localised 2c–2e C=O bonds, whereas $E[\text{QC}]$ favours retention of the aromatic tautomer. In *N*-heteroaromatic systems, the additional donor–acceptor stabilisation associated with lactam-type $n(\text{N}) \rightarrow \pi^*(\text{C}=\text{O})$ resonance and $n(\text{O}) \rightarrow \sigma^*(\text{C}–\text{N})$ interactions helps shift this balance towards the carbonylated tautomer.⁴⁴ In all-carbon benzenoid systems, the corresponding site-specific CH_2 -to-carbonyl hyperconjugation and $\sigma_{\text{C}–\text{C}}^*$ -based interactions are weaker, so carbonyl formation stabilises the keto forms but is not sufficient to overcome the aromatic preference. In the imine and methyldene series, the $\sigma^*(\text{C}–\text{N})$ -acceptor interactions increase only modestly upon tautomerisation, by +8.2 and +4.4 kcal mol^{-1} , respectively, compared with +30.9 kcal mol^{-1} for the carbonylated pair. These smaller stereoelectronic gains are therefore insufficient to offset the weaker interference stabilisation and larger quasi-classical penalties in **R7** and **R10**. Thus, the observed tautomeric preference is determined by the relative magnitude of these opposing contributions, rather than by aromaticity, carbonyl strength, or donor–acceptor effects alone.

3.3 Tautomeric preferences along a deamination pathway

To illustrate the predictive power of the additive model, we examine the successive deamination pathway of melamine (**9a**) to ammeline (**13a**), ammelide (**14b**), and cyanuric acid (**4b**).⁸⁹ Each intermediate admits hydroxy–carbonyl (lactim–lactam) tautomerism on the 1,3,5-triazine core, and in the literature these metabolites are often sketched interchangeably in either form. In practice, however, a single tautomer dominates for each species, and using the wrong depiction propagates confusion in reaction schemes. A consolidated scheme with the computed electronic

and free-energy differences for each tautomeric transformation and deamination step is provided in Fig. 7.

Within this framework, amino/imine substituents stabilise the aromatic tautomer, whereas hydroxy/carbonyl units favour the localised, non-aromatic alternative; in practice, two C=O/N–H units are usually required to overturn the aromatic preference. Applied to this series: melamine (three NH₂) strongly prefers the aromatic tautomer; ammeline (two NH₂ plus one OH) remains aromatic; ammelide (one NH₂ plus two C=O/N–H) crosses over to the non-aromatic, lactam-dominated tautomer; and cyanuric acid (three C=O/N–H) is firmly non-aromatic. These assignments align with the trends quantified above and rationalise which depictions are chemically faithful at each stage of the pathway.

Conclusions

In summary, we investigated heteroaromatic tautomeric equilibria and showed that their preferences arise primarily from the relative energetic balance between the exchanging functional groups and the aromatic ring framework. A simple ring/functional-group model rationalises the 2-hydroxypyridine/2-pyridone benchmark and extends across related series, where the introduction of C=O/N–H units shifts the equilibrium in an approximately additive manner, typically requiring two such units to overturn the aromatic preference. In contrast, amino/imine and methyl/methylidene exchanges reinforce the aromatic tautomer, consistent with the disfavouring of exocyclic C=N and C=CH₂ formation.

The SCGV/B/GPF-EP analysis provides a mechanistic interpretation of these trends. Specifically, *E*[QC] favours aromatic arrangements associated with a 6c–6e π -bonding pattern, whereas *E*[INT] favours the formation of localised, electron-precise 2c–2e bonds, particularly C=O units. This opposition provides an energy-decomposition expression of π -distortivity: interference, which encapsulates covalent contributions, stabilises bond localisation and bond-length alternation, while the quasi-classical term favours retention of the delocalised aromatic arrangement. The observed tautomeric preference therefore reflects the balance between these contributions, with carbonyl groups distinguished by a sufficiently large interference stabilisation to overcome the quasi-classical preference for aromaticity. The accompanying NBO analysis suggests a complementary stereoelectronic interpretation, where lactam-type carbonylated tautomers benefit from stronger $n(\text{N}) \rightarrow \pi^*(\text{C}=\text{O})$ resonance and associated $\sigma^*(\text{C}–\text{N})$ -acceptor interactions than the corresponding imine, methylidene, or all-carbon keto references.

The additive model is most directly applicable to six-membered *N*-heteroaromatic systems, but it also successfully explains energy differences between tautomeric forms in other aromatic series. In benzenoid phenols, for example, a similar regime is observed, although it leads to a different outcome: carbonyl formation also stabilises the keto forms, but the aromatic tautomer remains preferred even as the enol–keto gap decreases with increasing hydroxyl substitution, reflecting the larger penalty associated with disrupting an all-carbon π framework. The same balance between quasi-classical and interference contributions therefore applies,

but the relative magnitudes of these terms lead to a different tautomeric outcome.

Finally, the model rationalises the evolution of tautomeric preferences along the deamination pathway from melamine to cyanuric acid. Melamine and ammeline retain aromatic structures, whereas ammelide and cyanuric acid cross over to the carbonylated, non-aromatic forms once two or more C=O/N–H units are present. Although quantitative values may vary with improved correlation and solvation, this approach offers a simple and transferable basis for understanding and predicting tautomeric preferences across related molecular scaffolds.

Author contributions

L.A.: Methodology; Software; Formal analysis; Investigation; Data curation; Visualisation; Writing – original draft; Writing – review & editing. **F.F.:** Software; Validation; Resources; Visualisation; Supervision; Writing – original draft; Writing – review & editing; Funding acquisition. **P.M.E.:** Conceptualisation; Resources; Writing – review & editing; Supervision. **T.M.C.:** Methodology; Software; Writing – review & editing; Supervision; Project administration; Funding acquisition.

Conflicts of interest

There are no conflicts to declare.

Data availability

The data supporting this article have been included as part of the Supplementary Information.

Acknowledgements

The authors thank Aritra Roy and Esther R. S. Paz for computational assistance during the final stages of this work. F.F. acknowledges support from the Marie Skłodowska-Curie Actions under Grant Agreement Nos. 101299799 and 101225527; the Royal Society International Science Partnerships Fund through the International Collaboration Awards under Grant No. ICAO\R1\241112 and ICA Enhancement Funding under Grant No. ICAE\R1\261037; and the Royal Society of Chemistry Inclusion and Diversity Fund under Grant No. 280026109. T.M.C. thanks CNPq for financial support with the project Instituto Nacional de Ciência e Tecnologia em Ciências Moleculares (INCT-CiMol), Projeto CNPq 406804/2022-2. P.M.E. thanks CNPq, FAPERJ, and CAPES for financial support.

Notes and references

- 1 R. A. Sayle, *J. Comput. Aided. Mol. Des.*, 2010, **24**, 485–496.
- 2 M. K. Shukla and J. Leszczynski, *WIREs Comput. Mol. Sci.*, 2013, **3**, 637–649.
- 3 F. Wöhler, *Ann. Phys.*, 1828, **88**, 253–256.
- 4 C. A. Tsipis and P. A. Karipidis, *J. Am. Chem. Soc.*, 2003, **125**, 2307–2318.
- 5 P. Zuman and J. Michl, *Nature*, 1961, **192**, 655–657.
- 6 J. E. Dubois, M. El-Alaoui and J. Toullec, *J. Am. Chem. Soc.*, 1981, **103**, 5393–5401.
- 7 P. A. Nikitina, A. S. Peregodov, T. Y. Koldaeva, L. G. Kuz'mina,

- E. I. Adiulin, I. I. Tkach and V. P. Perevalov, *Tetrahedron*, 2015, **71**, 5217–5228.
- 8 J. Ortega-Castro, M. Adrover, J. Frau, A. Salvà, J. Donoso and F. Muñoz, *J. Phys. Chem. A*, 2010, **114**, 4634–4640.
- 9 A. S. Özen, P. Doruker and V. Aviyente, *J. Phys. Chem. A*, 2007, **111**, 13506–13514.
- 10 J. P. Lovie-Toon, C. M. Tram, B. L. Flynn and E. H. Krenske, *ACS Catal.*, 2017, **7**, 3466–3476.
- 11 P. Kozyra, A. Kaczor, Z. Karczmarzyk, W. Wysocki and M. Pitucha, *Struct. Chem.*, 2023, **34**, 1973–1984.
- 12 W. Tang, T. Yang, C. A. Morales-Rivera, X. Geng, V. K. Sri-rambhatla, X. Kang, V. P. Chauhan, S. Hong, Q. Tu, A. J. Florence, H. Mo, H. A. Calderon, C. Kisielowski, F. C. R. Hernandez, X. Zou, G. Mpourmpakis and J. D. Rimer, *Nat. Commun.*, 2023, **14**, 561.
- 13 W. Tang, T. Yang, Q. Tu, H. A. Calderon, F. C. Robles Hernandez, P. O'Reilly, Y. Jing, S. Park and J. D. Rimer, *Proc. Natl. Acad. Sci.*, 2025, **122**, e2426814122.
- 14 S. Carlotto, G. Casella, M. Sambi and M. Casarin, *Phys. Chem. Chem. Phys.*, 2021, **23**, 24661–24668.
- 15 Z. Li, Z. Wang, N. Chekshin, S. Qian, J. X. Qiao, P. T. Cheng, K.-S. Yeung, W. R. Ewing and J.-Q. Yu, *Science*, 2021, **372**, 1452–1457.
- 16 B. G. Janesko, H. C. Fisher, M. J. Bridle and J.-L. Montchamp, *J. Org. Chem.*, 2015, **80**, 10025–10032.
- 17 F.-U. Wehmeyer, Y. Li, A. Schlossarek, Z. Ke and R. Langer, *Dalt. Trans.*, 2025, **54**, 389–395.
- 18 S. Zhang, Q. Yang, L. Guo, Y. Zhang, L. Feng, L. Zhou, S. Yang, Q. Yao, G. Pescitelli and Z. Xie, *Chem. Commun.*, 2017, **53**, 10066–10069.
- 19 A. L. Hansen and L. E. Kay, *Proc. Natl. Acad. Sci.*, 2014, **111**, E1705–E1712.
- 20 Y. Tsuchiya, T. Tamura, M. Fujii and M. Ito, *J. Phys. Chem.*, 1988, **92**, 1760–1765.
- 21 J. Rejnek, M. Hanus, M. Kabeláč, F. Ryjáček and P. Hobza, *Phys. Chem. Chem. Phys.*, 2005, **7**, 2006–2017.
- 22 J. M. Bakker, R. K. Sinha, T. Besson, M. Brugnara, P. Tosi, J.-Y. Salpin and P. Maître, *J. Phys. Chem. A*, 2008, **112**, 12393–12400.
- 23 M. D. Topal and J. R. Fresco, *Nature*, 1976, **263**, 285–289.
- 24 W. Wang, H. W. Hellinga and L. S. Beese, *Proc. Natl. Acad. Sci.*, 2011, **108**, 17644–17648.
- 25 S. Chandorkar, S. Raghunathan, T. Jaganade and U. D. Priyakumar, *Int. J. Mol. Sci.*, 2021, **22**, 5411.
- 26 J. L. Siqueira-Neto, O.-R. Song, H. Oh, J.-H. Sohn, G. Yang, J. Nam, J. Jang, J. Cechetto, C. B. Lee, S. Moon, A. Genovesio, E. Chatelain, T. Christophe and L. H. Freitas-Junior, *PLoS Negl. Trop. Dis.*, 2010, **4**, e675.
- 27 M. J. Cook, A. R. Katritzky and P. Linda, *Advances in Heterocyclic Chemistry*, Academic Press, 1974, vol. 17, pp. 255–356.
- 28 T. M. Krygowski, H. Szatyłowicz, O. A. Stasyuk, J. Dominikowska and M. Palusiak, *Chem. Rev.*, 2014, **114**, 6383–6422.
- 29 G. Merino, M. Solà, I. Fernández, C. Foroutan-Nejad, P. Lazzeretti, G. Frenking, H. L. Anderson, D. Sundholm, F. P. Cossío, M. A. Petrukhina, J. Wu, J. I. Wu and A. Restrepo, *Chem. Sci.*, 2023, **14**, 5569–5576.
- 30 P. v. R. Schleyer, J. I. Wu, F. P. Cossío and I. Fernández, *Chem. Soc. Rev.*, 2014, **43**, 4909–4921.
- 31 A. Katritzky and J. Lagowski, *Advances in Heterocyclic Chemistry*, Academic Press, 1963, vol. 1, pp. 311–338.
- 32 A. T. Balaban, D. C. Oniciu and A. R. Katritzky, *Chem. Rev.*, 2004, **104**, 2777–2812.
- 33 J. Elguero, A. R. Katritzky and O. V. Denisko, *Advances in Heterocyclic Chemistry*, Academic Press, 2000, pp. 1–84.
- 34 A. R. Katritzky, *Chem. Heterocycl. Compd.*, 1972, **8**, 917–929.
- 35 M. J. Cook, A. R. Katritzky, P. Linda and R. D. Tack, *J. Chem. Soc. D Chem. Commun.*, 1971, 510.
- 36 M. J. Cook, A. R. Katritzky, P. Linda and R. D. Tack, *J. Chem. Soc. Perkin Trans. 2*, 1972, 1295.
- 37 H. B. Schlegel, P. Gund and E. M. Fluder, *J. Am. Chem. Soc.*, 1982, **104**, 5347–5351.
- 38 S. A. Hejazi, O. I. Osman, A. O. Alyoubi, S. G. Aziz and R. H. Hilal, *Int. J. Mol. Sci.*, 2016, **17**, 1893.
- 39 F. Weinhold, C. Landis and E. Glendening, *Int. Rev. Phys. Chem.*, 2016, **35**, 399–440.
- 40 A. S. Gazizov, A. V. Smolobochkin, T. S. Rizbayeva, S. Z. Vatsadze, A. R. Burilov, O. G. Sinyashin and I. V. Alabugin, *J. Org. Chem.*, 2023, **88**, 6868–6877.
- 41 T. Lukmanov, S. P. Ivanov, E. M. Khamitov and S. L. Khursan, *Comput. Theor. Chem.*, 2013, **1023**, 38–45.
- 42 J. I. Seeman and D. J. Tantillo, *Chem. Sci.*, 2022, **13**, 11461–11486.
- 43 I. V. Alabugin, L. Kuhn, N. V. Krivoshchapov, P. Mehaffy and M. G. Medvedev, *Chem. Soc. Rev.*, 2021, **50**, 10212–10252.
- 44 I. V. Alabugin, L. Kuhn, M. G. Medvedev, N. V. Krivoshchapov, V. A. Vil', I. A. Yaremenko, P. Mehaffy, M. Yarie, A. O. Terent'ev and M. A. Zolfigol, *Chem. Soc. Rev.*, 2021, **50**, 10253–10345.
- 45 T. L. P. Galvão, I. M. Rocha, M. D. M. C. Ribeiro da Silva and M. A. V. Ribeiro da Silva, *J. Phys. Chem. A*, 2013, **117**, 5826–5836.
- 46 H. Meyer, *Monatsh. Chem.*, 1905, **26**, 1303–1310.
- 47 L. D. Hatherley, R. D. Brown, P. D. Godfrey, A. P. Pierlot, W. Caminati, D. Damiani, S. Melandri and L. B. Favero, *J. Phys. Chem.*, 1993, **97**, 46–51.
- 48 A. L. Sobolewski, *Chem. Phys. Lett.*, 1993, **211**, 293–299.
- 49 P. Beak and F. S. Fry, *J. Am. Chem. Soc.*, 1973, **95**, 1700–1702.
- 50 M. J. Scanlan and I. H. Hillier, *Chem. Phys. Lett.*, 1984, **107**, 330–332.
- 51 J. Smets and G. Maes, *Chem. Phys. Lett.*, 1991, **187**, 532–536.
- 52 M. Takasuka, T. Saito and H. Nakai, *Vib. Spectrosc.*, 1996, **13**, 65–74.
- 53 D. H. Aue, L. D. Betowski, W. R. Davidson, M. T. Bowers, P. Beak and J. Lee, *J. Am. Chem. Soc.*, 1979, **101**, 1361–1368.
- 54 B. R. Penfold, *Acta Crystallogr.*, 1953, **6**, 591–600.
- 55 M. R. Nimlos, D. F. Kelley and E. R. Bernstein, *J. Phys. Chem.*, 1989, **93**, 643–651.
- 56 M. J. Cook, A. R. Katritzky, M. Taagepera, T. D. Singh and

- R. W. Taft, *J. Am. Chem. Soc.*, 1976, **98**, 6048–6049.
- 57 A. R. Katritzky, *Advances in Heterocyclic Chemistry*, Academic press, 1997, vol. 70.
- 58 E. W. Hughes, *J. Am. Chem. Soc.*, 1941, **63**, 1737–1752.
- 59 H. W. Yang and B. M. Craven, *Acta Crystallogr. Sect. B Struct. Sci.*, 1998, **54**, 912–920.
- 60 V. Vaz da Cruz, R. Büchner, M. Fondell, A. Pietzsch, S. Eckert and A. Föhlisch, *J. Phys. Chem. Lett.*, 2022, **13**, 2459–2466.
- 61 M. J. Nowak, L. Lapinski, J. Fulara, A. Les and L. Adamowicz, *J. Phys. Chem.*, 1992, **96**, 1562–1569.
- 62 W. S. Hamama, M. Waly, I. El-Hawary and H. H. Zoorob, *Synth. Commun.*, 2014, **44**, 1730–1759.
- 63 S. Moreno, M. Brunner, I. Delazer, D. Rieder, A. Lusser and R. Micura, *RSC Chem. Biol.*, 2022, **3**, 447–455.
- 64 P. Rubovič, A. Pysanenko, J. Lengyel, D. Nachtigallová and M. Fárnik, *J. Phys. Chem. A*, 2016, **120**, 4720–4730.
- 65 F. Neese, *WIREs Comput. Molec. Sci.*, 2012, **2**, 73–78.
- 66 C. Riplinger and F. Neese, *J. Chem. Phys.*, 2013, **138**, 034106.
- 67 M. Friede, S. Ehlert, S. Grimme and J.-M. Mewes, *J. Chem. Theory Comput.*, 2023, **19**, 8097–8107.
- 68 J. Zheng, X. Xu and D. G. Truhlar, *Theor. Chem. Acc.*, 2011, **128**, 295–305.
- 69 J. Li and R. McWeeny, *Int. J. Quantum Chem.*, 2002, **89**, 208–216.
- 70 G. M. Barca, C. Bertoni, L. Carrington, D. Datta, N. De Silva, J. E. Deustua, D. G. Fedorov, J. R. Gour, A. O. Gunina, E. Guidez *et al.*, *J. Chem. Phys.*, 2020, **152**, 154102.
- 71 T. M. Cardozo and M. A. C. Nascimento, *J. Chem. Phys.*, 2009, **130**, 104102.
- 72 K. Ruedenberg, *Rev. Mod. Phys.*, 1962, **34**, 326–376.
- 73 R. McWeeny, *Proc. R. Soc. London. Ser. A. Math. Phys. Sci.*, 1959, **253**, 242–259.
- 74 T. M. Cardozo, F. Fantuzzi and M. A. C. Nascimento, *Phys. Chem. Chem. Phys.*, 2014, **16**, 11024–30.
- 75 D. L. Cooper, J. Gerratt and M. Raimondi, *Nature*, 1986, **323**, 699–701.
- 76 P. C. Hiberty, D. Danovich, A. Shurki and S. Shaik, *J. Am. Chem. Soc.*, 1995, **117**, 7760–7768.
- 77 M. W. Wong, K. B. Wiberg and M. J. Frisch, *J. Am. Chem. Soc.*, 1992, **114**, 1645–1652.
- 78 V. Barone and M. Cossi, *J. Phys. Chem. A*, 1998, **102**, 1995–2001.
- 79 A. V. Marenich, C. J. Cramer and D. G. Truhlar, *J. Phys. Chem. A*, 2009, **113**, 6378–6396.
- 80 J. Frank and A. R. Katritzky, *J. Chem. Soc., Perkin Trans. 2*, 1976, 1428–1431.
- 81 M. Kuzuya, A. Noguchi and T. Okuda, *J. Chem. Soc., Perkin Trans. 2*, 1985, 1423–1427.
- 82 M. Hatanaka, *Tetrahedron Lett.*, 2013, **54**, 1452–1455.
- 83 S. Shaik, A. Shurki, D. Danovich and P. C. Hiberty, *J. Am. Chem. Soc.*, 1996, **118**, 666–671.
- 84 S. Shaik, A. Shurki, D. Danovich and P. C. Hiberty, *Chem. Rev.*, 2001, **101**, 1501–1540.
- 85 R. S. Berry, *J. Chem. Phys.*, 1961, **35**, 2253–2254.
- 86 Y. Haas and S. Zilberg, *J. Am. Chem. Soc.*, 1995, **117**, 5387–5388.
- 87 F. Fantuzzi and M. A. C. Nascimento, *J. Chem. Theory Comput.*, 2014, **10**, 2322–2332.
- 88 I. V. Alabugin, K. M. Gilmore and P. W. Peterson, *WIREs Comput. Mol. Sci.*, 2011, **1**, 109–141.
- 89 M. H. Almatarneh, A. A.-A. A. Abu-Saleh and I. A. Elayan, *Comput. Theor. Chem.*, 2017, **1117**, 92–99.

Data availability

The data supporting this article have been included as part of the Supplementary Information.